Available online at www.sciencedirect.com
ScienceDirect
Journal of Multivariate Analysis 99 (2008) 1418–1437
Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

A bivariate Lévy process with negative binomial and gamma marginals

Tomasz J. Kozubowski[a,*], Anna K. Panorska[a], Krzysztof Podgórski[b]

[a] Department of Mathematics and Statistics, University of Nevada, Mail Stop 84, 89557-0045 Reno, NV, United States
[b] Lund University, Sweden

Available online 17 February 2008

Abstract

The joint distribution of X and N, where N has a geometric distribution and X is the sum of N IID exponential variables (independent of N), is infinitely divisible. This leads to a bivariate Lévy process $\{(X(t), N(t)), t \geq 0\}$, whose coordinates are correlated negative binomial and gamma processes. We derive basic properties of this process, including its covariance structure, representations, and stochastic self-similarity. We examine the joint distribution of $(X(t), N(t))$ at a fixed time t, along with the marginal and conditional distributions, joint integral transforms, moments, infinite divisibility, and stability with respect to random summation. We also discuss maximum likelihood estimation and simulation for this model.

AMS 2000 subject classifications: 60E05; 60E07; 60F05; 60G18; 60G50; 60G51; 62H05; 62H12

Keywords: Discrete Lévy process; Gamma process; Gamma Poisson process; Infinite divisibility; Maximum likelihood estimation; Negative binomial process; Operational time; Random summation; Random time transformation; Stability; Subordination; Self-similarity

* Corresponding author.
E-mail address: tkozubow@unr.edu (T.J. Kozubowski).

0047-259X/$ - see front matter
doi:10.1016/j.jmva.2008.02.029

1. Introduction

Kozubows and Panorska [11] studied a bivariate distribution with exponential and geometric marginals (BEG), denoted by $\mathcal{BEG}(\beta, p)$ and defined via the stochastic representation

$$(X, N) \stackrel{d}{=} \left(\sum_{i=1}^{N} X_i, N\right). \tag{1}$$

Here, the $\{X_i\}$ are IID exponential variables with the PDF

$$f(x) = \beta e^{-\beta x}, \quad x > 0, \tag{2}$$

and N is a geometric random variable with the PDF

$$h(n) = \mathbb{P}(N = n) = p(1-p)^{n-1}, \quad n = 1, 2, \ldots, \tag{3}$$

independent of the $\{X_i\}$. They derived basic properties of these models, including marginal and conditional distributions, joint integral transforms, infinite divisibility, stability with respect to geometric summation, and maximum likelihood estimation of the parameters. Applications of the BEG model range from finance and actuarial science to hydrology and climate. In the financial application presented in [11], the process $\{X_i\}$ represented log-returns of the exchange rates between two currencies for their growth/decline periods. The BEG model then provided the joint distribution for the cumulative log-returns and the length of their growth/decline periods. The fit of the BEG model to this data was quite remarkable in terms of marginal as well and bivariate distributions. Similarly, BEG model yields itself useful in actuarial problems of modeling total size of the claims exceeding a certain threshold when the number of claims is geometric.

In hydrology and climate research, the problem of estimating total stream flow, precipitation, Palmer Drought Severity Index, or Pacific Decadal Oscillation index exceeding a threshold (e.g. long term mean or high percentile) is of primary importance to water resource managers and safe engineering design needing reasonable flood, drought, or water storage estimates (see, e.g., [2,6,14,17,18]). These problems are usually studied in terms of *episodes*, represented as a random vector of magnitude and duration, where the duration N is the number of time intervals (e.g. years) the process remains continuously above (or below) a reference level, while the magnitude X is the sum of all process values for a given duration. Here, the BEG model provides a stochastic framework for the joint distribution of (X, N), and has been successfully applied to stream flow analysis in [3].

While the BEG model proved to be quite useful, in many applications a more general model is needed. One important extension arises as the sum of n independent BEG random vectors, whose distribution is the same as the marginal distribution at $t = n$ of a bivariate Lévy process $\{(X(t), N(t)), t \geq 0\}$, where $(X(1), N(1))$ is given by (1). This process, which was briefly mentioned in [11], admits a stochastic representation

$$\{(X(t), N(t)),\ t \geq 0\} \stackrel{d}{=} \left\{\left(\sum_{i=1}^{NB(t)} X_i + G(t), NB(t) + t\right),\ t \geq 0\right\}, \tag{4}$$

where the $\{X_i\}$ are as before, $\{G(t),\ t \geq 0\}$ is a gamma Lévy process starting at zero, based on the exponential distribution (2), and $\{NB(t), t \geq 0\}$ is a negative binomial (NB) Lévy process starting at zero, with the ChF

$$\mathbb{E}e^{isNB(t)} = \left(\frac{p}{1-(1-p)e^{is}}\right)^t, \quad s \in \mathbb{R}, \tag{5}$$

studied by Kozubowski and Podgórski [12,13]. Three other related Lévy processes are obtained by replacing either t or $G(t)$ (or both) on the right-hand-side of (4) by zero. These continuous-time models have high potential for use in stochastic modeling involving negative binomial sums of independent random quantities. For example, in insurance applications the above process (with deleted t and $G(t)$) would describe the joint behavior of severity and frequency of claims over time, assuming the claims are exponentially distributed and the counting process of the number of claims is negative binomial.

In this paper we focus on the process obtained by deleting t from (4), so that for each $t > 0$ the coordinate $N(t)$ has the negative binomial distribution (5). One can think of this process as a bivariate Lévy process $\{(X(t), NB(t)), t \geq 0\}$ whose marginal distribution at $t = 1$ is the same as that of $(X, N-1)$, with (X, N) defined via (1). Using the results from [11], we find that the characteristic function (ChF) of this Lévy process is

$$\mathbb{E}e^{i\{rX(t)+sNB(t)\}} = \left(\frac{p\beta}{\beta - ir - \beta(1-p)e^{is}}\right)^t, \quad r, s \in \mathbb{R}. \tag{6}$$

Clearly, $\{X(t), t \geq 0\}$ is a gamma Lévy process starting at zero with the ChF

$$\mathbb{E}e^{irX(t)} = \left(\frac{p\beta}{p\beta - ir}\right)^t, \quad r \in \mathbb{R}. \tag{7}$$

Since the one-dimensional coordinates have gamma and negative binomial distributions, we call this process a BGNB process (**B**ivariate Lévy process with **g**amma and **n**egative **b**inomial marginals) and if we want reference to the parameters, we call it the $\mathcal{BGNB}(\beta, p)$ Lévy motion. Both gamma and negative binomial Lévy process play an increasingly important role in applications, where they are often used as subordinators of Gaussian or Poisson processes (see, e.g., [5,10,12,13]). Combining these two processes into a single bivariate correlated process leads to a new distribution with high potential use in stochastic modeling.

Our paper is organized as follows. In Section 2 we present various representations of the bivariate Lévy process, along with its invariance properties related to subordination. In Section 3 we derive basic properties of its bivariate marginal distributions, including joint densities, integral transforms, distribution and survival functions, moments and the covariance structure, and certain conditional distributions. Sections 4 and 5 are devoted to parameter estimation and simulation, respectively. In Section 6 we briefly discuss three other processes connected with the representation (4). Finally, proofs and auxiliary results are collected in Section 7.

2. Representations and stability properties

2.1. Stochastic representations

We start with two stochastic representations of BGNB process. The first one is related to subordination of the gamma process.

Proposition 2.1. *Let $\{NB(t), t \geq 0\}$ be a negative binomial (NB) Lévy process started at zero defined by (5) and let $\{G(t), t \geq 0\}$ be an independent gamma Lévy process, where $G(1)$ has the* *exponential distribution with PDF (2). Then the BGNB process given by the ChF (6) admits the representation*

$$\{(X(t), NB(t)),\ t \geq 0\} \stackrel{d}{=} \{(G(NB(t)+t), NB(t)),\ t \geq 0\}. \tag{8}$$

This follows from the *stochastic self-similarity* property: a gamma process subordinated to a NB process with drift, $NB(t) + t$, is again a gamma process (see [10], [13]). To see (8), consider the marginal distributions of these (Lévy) processes at $t = 1$. Here, the distribution corresponding to the left-hand-side of (8) is the same as that of $(X, N - 1)$, where (X, N) is given by (1). On the other hand, we have

$$G(NB(1)+1) \stackrel{d}{=} G(N) = (G(1) - G(0)) + \cdots + (G(N) - G(N-1)), \tag{9}$$

where the terms $G(j) - G(j-1)$, $j = 1, \ldots, N$, are IID exponential variables with density (2), so that the right-hand-side of (8) has the same distribution as the left-hand-side. Thus, starting with two independent Lévy processes, a NB process with parameter $p \in (0, 1)$ defined by (5) and a gamma process with scale $\beta > 0$ defined by

$$\mathbb{E}e^{irG(t)} = \left(\frac{\beta}{\beta - ir}\right)^t, \quad r \in \mathbb{R}, \tag{10}$$

we can construct a bivariate Lévy process with the ChF (6) via (8).

Remark 2.1. In the notation of Proposition 2.1, the representation (8) admits an alternative formulation in the spirit of (4),

$$\{(X(t), NB(t)),\ t \geq 0\} \stackrel{d}{=} \left\{\left(\sum_{i=1}^{NB(t)} X_i + G(t), NB(t)\right),\ t \geq 0\right\}, \tag{11}$$

where the $\{X_i\}$ are IID exponential variables with PDF (2).

Our second representation is related to subordination of a Poisson process. Recall that a NB process given by (5) admits a representation $NB(t) \stackrel{d}{=} N(G(t))$, where $\{G(t), t \geq 0\}$ is a gamma process with parameter $\gamma > 0$ (so that $G(1)$ has an exponential distribution with mean $1/\gamma$) and $\{N(t), t \geq 0\}$ is an independent Poisson process with parameter $\lambda = \gamma(1-p)/p$, see, e.g., [12]. The result below extends this to BGNB process.

Proposition 2.2. *Let $\{G(t), t \geq 0\}$ be a gamma Lévy process with parameter $\gamma = p\beta$ given by the ChF (7) and let $\{N(t), t \geq 0\}$ be an independent Poisson process with parameter $\lambda = \beta(1 - p)$. Then the BGNB process given by the ChF (6) admits the representation*

$$\{(X(t), NB(t)),\ t \geq 0\} \stackrel{d}{=} \{(G(t), N(G(t))),\ t \geq 0\}. \tag{12}$$

From the above proposition one can derive the asymptotics of a BGNB process for small p, which follow directly from the almost sure convergence of $N(u)/u$ to one as $u \to \infty$.

Corollary 2.1. *Let $G(t)$ and $N(t)$ be independent standard gamma and Poisson Lévy processes, respectively. Then the process*

$$\mathbf{Y}_p(t) = \left(\frac{G(t)}{p\beta}, N\left(\frac{\bar{p}G(t)}{p}\right)\right),$$

where $\bar{p} = 1 - p$*, is a BGNB process. Moreover, for each* $t > 0$*, with probability one*

$$\lim_{p\to 0} p\mathbf{Y}_p(t) = (G(t)/\beta, G(t)).$$

2.2. Stochastic self-similarity and invariance properties

Here we study invariance properties of BGNB processes with respect to time deformation. Let $\{(X(t), NB(t)), t \geq 0\}$ be a BGNB process with parameters $p \in (0, 1)$ and $\beta > 0$, and let $\{\widetilde{NB}_q(t), t \geq 0\}$ be an independent NB process with parameter $q \in (0, 1)$. Then, as discussed in [10] and [13], the processes $X(t)$ and $NB(t)$ subordinated to $\widetilde{NB}_q(t) + t$ are again gamma and NB processes, respectively. It is not surprising that our bivariate process enjoys this invariance property as well.

Proposition 2.3. *If* $\{(X(t), NB(t)), t \geq 0\}$ *is a BGNB process with parameters* $p \in (0, 1)$ *and* $\beta > 0$*, then in the above setting the subordinated process*

$$\{(X(\widetilde{NB}_q(t) + t), NB(\widetilde{NB}_q(t) + t)), t \geq 0\} \tag{13}$$

is a BGNB process with parameters

$$p^* = \frac{pq}{1 - p + pq} \quad \textit{and} \quad \beta^* = \beta(1 - p + pq). \tag{14}$$

Remark 2.2. The fact that the subordinated process $NB(\widetilde{NB}_q(t) + t)$ is again a NB process with parameter p^*, noted by Kozubowski and Podgórski [13], generalizes the well-known stability property of geometric distribution with respect to geometric convolutions. The latter states that if X_i are IID geometric variables with parameter p given by the PDF (3) and N_q is another geometric variable with parameter q, independent of the $\{X_i\}$, then the sum $\sum_{j=1}^{N_q} X_j$ has also geometric distribution with parameter pq.

Remark 2.3. This invariance of the gamma process can be stated as follows:

$$\{X(T_c(t))\ t \geq 0\} \stackrel{d}{=} \{c^H X(t), t \geq 0\}, \quad c \geq 1, \tag{15}$$

where $c = 1/q$, $H = 1$, and $T_c(t) = \widetilde{NB}_{1/c}(t) + t$. Note that $\mathcal{T} = \{T_c(t), t \geq 0\}$, $c \geq 1$, is a family of non-negative and non-decreasing stochastic processes whose expectations are linear in t, $\mathbb{E}T_c(t) = ct$. In analogy with classical self-similarity, Kozubowski et al. [10] call processes $X(t)$ satisfying (15) *stochastically self-similar* (SSS) with index H with respect to family $\mathcal{T}$. Other examples of SSS processes (with respect to the same family of NB time changes), discussed in [10], include *fractional Laplace motion* (defined as a fractional Brownian motion subordinated to a gamma process) and, more generally, any other self-similar process subordinated to a gamma process.

2.3. Series representations

The classical series representations of Lévy motions are a convenient way to summarize the sample path properties of pure jump processes, where the distributions of values of jumps and their position are provided (see [4,8]). We have the following result for the BGNB process.

Proposition 2.4. *Let* $\mathbf{Y}(t)$ *be a* $\mathcal{BGNB}(\beta, p)$ *Lévy process and let* $G(t)$ *be a standard gamma process with scale* β. *Further, let* $\{\Gamma_i\}$ *be a sequence of arrivals of a standard Poisson process,* $\{E_i\}$ *be a sequence of IID standard exponential random variables,* $\{N_i\}$ *be a sequence of independent standard Poisson processes, and* $\{V_i\}$ *be a sequence of independent uniform random variables, where all the sequences are mutually independent. Then* $\mathbf{Y}(t)$ *can be represented in distribution as the sum of two independent Lévy processes,*

$$\mathbf{Y}(t) = \mathbf{A}\left(\mathbf{Y}_1(t) + \mathbf{Y}_2(t)\right).$$

Here,

$$\mathbf{A} = \begin{bmatrix} 1/\beta & 0 \\ 0 & 1 \end{bmatrix},$$

$\mathbf{Y}_1(t) = (G(t), 0)$, *and*

$$\mathbf{Y}_2(t) = \sum_{i=1}^{\infty} \mathbf{J}_i \mathbf{1}_{[\Gamma_i,\infty)}(\lambda t),$$

where $\lambda = -\log p$ *and the jumps are*

$$\mathbf{J}_i = \left(\frac{E_i}{p^{V_i}}, N_i\left((1 - p^{V_i})\frac{E_i}{p^{V_i}}\right)\right).$$

In particular, conditionally on $V_i = v$, *the vector* $\mathbf{J}_i$ *has the distribution given by the value of* $\mathcal{BGNB}(1, p^v)$ *Lévy motion at time one.*

3. Distributions of the bivariate process

3.1. Marginal distributions

Here we consider marginal bivariate distributions of the BGNB process defined by the ChF (6) at a fixed time t. To derive the PDF, we use the representation (8) given in Proposition 2.1. Clearly, the random variable $NB(t)$ has a negative binomial distribution supported on non-negative integers, given by the ChF (5) and the PDF

$$f_{NB(t)}(k) = \mathbb{P}(NB(t) = k) = \frac{\Gamma(k+t)}{k!\Gamma(t)} p^t (1-p)^k, \quad k = 0, 1, 2, \ldots. \tag{16}$$

Moreover, the distribution of $X(t)$ is gamma with shape parameter t and scale $p\beta$, given by the ChF (7). On the other hand, the conditional distribution of $X(t)$ given $NB(t) = k, k = 0, 1, 2\ldots,$ is gamma with shape parameter $t + k$ and scale β, so that the conditional PDF is

$$f_{X(t)|NB(t)=k}(x) = \frac{\beta^{k+t}}{\Gamma(t+k)} x^{k+t-1} \mathrm{e}^{-\beta x}, \quad x > 0. \tag{17}$$

Consequently, the joint PDF of $X(t)$ and $NB(t)$ is

$$g(x,k) = \frac{\beta^{t+k}}{k!\Gamma(t)} x^{k+t-1} \mathrm{e}^{-\beta x} p^t (1-p)^k, \quad x > 0, k = 0, 1, 2, \ldots. \tag{18}$$

Note that in case $t = 1$ this simplifies to the PDF of the (shifted) BEG model discussed in [11]. It is now easy to show that the conditional PDF of $NB(t)$ given $X(t) = x$ is given by

$$f_{NB(t)|X(t)=x}(k) = \frac{[\beta(1-p)x]^k \mathrm{e}^{-\beta(1-p)x}}{k!}, \quad k = 0, 1, 2, \dots, \tag{19}$$

which is a Poisson distribution with mean $\beta(1-p)x$. Since the marginal distributions of $X(t)$ and $NB(t)$ are gamma and negative binomial, respectively, we shall refer to this distribution as the BGNB distribution (**b**ivariate with **g**amma and **n**egative **b**inomial marginals).

Definition 3.1. A random vector (X, Y) with the PDF (18) is said to have the BGNB distribution with parameters $t > 0$, $\beta > 0$ and $p \in (0, 1)$. This distribution is denoted by $\mathcal{BGNB}(t, \beta, p)$.

Remark 3.1. It follows from (11) that a BGNB random vector with parameters $t > 0$, $\beta > 0$ and $p \in (0, 1)$ admits a stochastic representation

$$(X, Y) \stackrel{d}{=} \left(\sum_{i=1}^{T} E_i + G, T \right), \tag{20}$$

where all the variables on the right-hand-side of (20) are mutually independent, the $\{E_i\}$ are IID exponential variables with PDF (2), G has a gamma distribution with shape parameter t and scale β, and T is a NB variable with the PDF (16).

The cumulative distribution function (CDF) and the survival function (SF) are easily derived from the PDF (18).

Proposition 3.1. *If* $(X, Y) \sim \mathcal{BGNB}(t, \beta, p)$ *then for any* $x, y \geq 0$ *we have*

$$P(X \leq x, Y \leq y) = \frac{p^t}{\Gamma(t)} \sum_{j=0}^{[y]} \frac{(1-p)^j}{j!} \{\Gamma(j+t) - \Gamma(j+t, \beta x)\} \tag{21}$$

and

$$P(X > x, Y > y) = \frac{\Gamma(t, p\beta x)}{\Gamma(t)} - \frac{p^t}{\Gamma(t)} \sum_{j=0}^{[y]} \frac{(1-p)^j}{j!} \Gamma(j+t, \beta x), \tag{22}$$

where $[y]$ *is the integer part of* y *and*

$$\Gamma(\alpha, x) = \int_x^\infty w^{\alpha-1} \mathrm{e}^{-w} \mathrm{d}w \tag{23}$$

is the incomplete gamma function.

We skip routine calculations leading to the joint moments of BGNB RVs.

Proposition 3.2. *If* $(X, Y) \sim \mathcal{BGNB}(t, \beta, p)$*, then for any* $\eta, \gamma \geq 0$ *we have*

$$\mathbb{E}X^\eta Y^\gamma = \frac{\Gamma(\eta + t)}{\Gamma(t)} \left(\frac{1}{p\beta} \right)^\eta \mu_\gamma, \tag{24}$$

where $\mu_\gamma = \mathbb{E}W^\gamma$ *with* W *having a NB distribution with parameters* $\eta + t > 0$ *and* $p \in (0, 1)$.

Special cases of this result lead to the mean vector and the covariance matrix of (X, Y), presented below without proof.

Corollary 3.1. *If* $(X, Y) \sim \mathcal{BGNB}(t, \beta, p)$, *then* $\mathbb{E}X = t(\beta p)^{-1}$, $\mathbb{E}Y = t(1-p)p^{-1}$, *the covariance matrix of* (X, Y) *is*

$$\Sigma = t \cdot \begin{bmatrix} \frac{1}{\beta^2 p^2} & \frac{1-p}{\beta p^2} \\ \frac{1-p}{\beta p^2} & \frac{1-p}{p^2} \end{bmatrix}, \tag{25}$$

and the correlation coefficient of X *and* Y *is* $\rho = \sqrt{1-p}$.

The moments of the coordinates of $\mathcal{BGNB}(\beta, p)$ Lévy motion follow from the above. Since the process has independent increments its covariance function is easily derived and given in the following result.

Corollary 3.2. *If* $\mathbf{Y}(t)$ *is a* $\mathcal{BGNB}(\beta, p)$ *Lévy motion then*

$$\mathrm{Cov}(\mathbf{Y}(t), \mathbf{Y}(s)) = \frac{t \wedge s}{p^2} \begin{bmatrix} \frac{1}{\beta^2} & \frac{1-p}{\beta} \\ \frac{1-p}{\beta} & 1-p \end{bmatrix}.$$

3.2. *Certain conditional distributions*

Here we present certain conditional distributions of $(X, Y) \sim \mathcal{BGNB}(t, \beta, p)$, which are useful in the practical implementation and goodness-of-fit analysis of these laws.

3.2.1. *The conditional distribution of* (X, Y) *given* $Y > n$

For real $x > 0$ and integers $m, n \geq 0$ we have

$$\mathbb{P}(X > x, Y > m | Y > n) = \frac{1}{c_n} \left\{ \frac{\Gamma(t, p\beta x)}{\Gamma(t)} - \frac{p^t}{\Gamma(t)} \sum_{j=0}^{\max(m,n)} \frac{(1-p)^j}{j!} \Gamma(j+t, \beta x) \right\},$$

where

$$c_n = \mathbb{P}(Y > n) = 1 - \sum_{j=0}^{n} \frac{\Gamma(j+t)}{j!\Gamma(t)} p^t (1-p)^j.$$

When $0 \leq m \leq n$ the above represents the SF of X given $Y > n$, $\mathbb{P}(X > x | Y > n)$. The corresponding PDF is

$$f_{X|Y>n}(x) = \frac{1}{c_n} \frac{(p\beta)^t}{\Gamma(t)} x^{t-1} \mathrm{e}^{-p\beta x} \left\{ 1 - \mathrm{e}^{-(1-p)\beta x} \sum_{j=0}^{n} \frac{((1-p)\beta x)^j}{j!} \right\}, \quad x > 0.$$

3.2.2. *The conditional distribution of* (X, Y) *given* $X > u$

Similarly, for integer $n \geq 0$ and real $x, u > 0$ we have

$$\mathbb{P}(X > x, Y > n | X > u) = \frac{\Gamma(t, p\beta \max(x,u))}{\Gamma(t, p\beta u)} - \frac{p^t}{\Gamma(t, p\beta u)} \sum_{j=0}^{n} \frac{(1-p)^j}{j!} \Gamma(j+t, \beta \max(x,u)).$$

In case $u \geq x > 0$, the above represents the SF of Y given $X > u$, with the corresponding PDF

$$\mathbb{P}(Y = n | X > u) = \frac{p^t (1-p)^n}{n!} \frac{\Gamma(n+t, \beta u)}{\Gamma(t, p\beta u)}, \quad n = 0, 1, \ldots.$$

4. Estimation

In this section we consider maximum likelihood estimation of the BGNB parameters. Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be a random sample from a $\mathcal{BGNB}(\alpha, \beta, p)$ distribution. The log-likelihood function takes the form

$$L(\alpha, \beta, p) = n \left\{ \overline{Y}_n \log[\beta(1-p)] + \alpha[\log(\beta p) + \overline{(\log X)}_n] - \log \Gamma(\alpha) - \beta \overline{X}_n + C \right\}, \tag{26}$$

where $\overline{X}_n$ and $\overline{Y}_n$ are sample means of the $\{X_i\}$ and the $\{Y_i\}$ respectively, $\overline{(\log X)}_n$ is the sample mean of the quantities $\log X_i$, and

$$C = \frac{1}{n} \sum_{i=1}^{n} \log \frac{X_i^{Y_i - 1}}{Y_i!}.$$

It is easy to see that for a fixed $\alpha > 0$ the function (26) is maximized by the pair

$$\hat{\beta}_n = \hat{\beta}_n(\alpha) = \frac{\alpha + \overline{Y}_n}{\overline{X}_n}, \quad \hat{p}_n = \hat{p}_n(\alpha) = \frac{\alpha}{\alpha + \overline{Y}_n}, \tag{27}$$

which are the maximum likelihood estimators (MLEs) of β and p when α is known. When we substitute these values back into (26) we end up with the problem of maximizing the function

$$R(\alpha) = -\alpha \log \overline{X}_n - \log \Gamma(\alpha) + \alpha \overline{(\log X)}_n - \alpha + \alpha \log \alpha \tag{28}$$

with respect to $\alpha > 0$. The following result shows that this problem has a unique solution with probability one.

Proposition 4.1. *Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be IID observations from a $\mathcal{BGNB}(\alpha, \beta, p)$ distribution. If not all $\{X_j\}$ are equal, then there exist unique MLEs of α, β, and p, denoted by $\hat{\alpha}_n$, $\hat{\beta}_n$ and $\hat{p}_n$, respectively. Moreover, $\hat{\alpha}_n$ is a unique solution of the equation*

$$\overline{(\log X)}_n - \log \overline{X}_n + \log \alpha - \psi(\alpha) = 0, \tag{29}$$

where ψ is the Digamma function, while $\hat{\beta}_n$ and $\hat{p}_n$ are given by (27) with α replaced by $\hat{\alpha}_n$.

Standard large sample theory shows that the above MLEs are consistent, asymptotically normal, and efficient. Routine calculations show that the Fisher information matrix

$$I(\alpha, \beta, p) = \left[-E \left(\frac{\partial^2}{\partial \theta_i \partial \theta_j} \log g_{\alpha,\beta,p}(X, Y) \right) \right]_{i,j=1}^{3} \tag{30}$$

corresponding to the $\mathcal{BGNB}(\alpha, \beta, p)$ distribution with the vector-parameter $\theta = (\theta_1, \theta_2, \theta_3)' = (\alpha, \beta, p)'$ and density $g_{\alpha,\beta,p}$ is given by

$$I(\alpha, \beta, p) = \begin{bmatrix} \psi'(\alpha) & -\dfrac{1}{\beta} & -\dfrac{1}{p} \\ -\dfrac{1}{\beta} & \dfrac{\alpha}{p\beta^2} & 0 \\ -\dfrac{1}{p} & 0 & \dfrac{\alpha}{p^2(1-p)} \end{bmatrix}. \tag{31}$$

Here, $\psi'(\alpha)$ is the Trigamma function

$$\psi'(\alpha) = \sum_{n=0}^{\infty} \frac{1}{(\alpha + n)^2},$$

i.e. the derivative of the Digamma function $\psi(\alpha) = \frac{d}{d\alpha} \log \Gamma(\alpha)$.

Proposition 4.2. *The maximum likelihood estimators of α, β and p based on a random sample $(X_1, Y_1), \ldots, (X_n, Y_n)$ from a $\mathcal{BGNB}(\alpha, \beta, p)$ distribution are*

(i) *consistent;*

(ii) *asymptotically normal, that is $\sqrt{n}[(\hat{\alpha}_n, \hat{\beta}_n, \hat{p}_n) - (\alpha, \beta, p)]$ converges in distribution to a trivariate normal distribution with the (vector) mean zero and the covariance matrix*

$$\Sigma_{MLE} = \frac{1}{\alpha C} \cdot \begin{bmatrix} \alpha^2 & \alpha\beta p & \alpha p(1-p) \\ \alpha\beta p & \beta^2 p(p+C) & \beta p^2(1-p) \\ \alpha p(1-p) & \beta p^2(1-p) & p^2(1-p)(1-p+C) \end{bmatrix}, \tag{32}$$

where $C = \alpha\psi'(\alpha) - 1$ and $\psi'(\alpha)$ is the Trigamma function;

(iii) *asymptotically efficient — the asymptotic covariance matrix (32) coincides with the inverse of the Fisher information matrix (31).*

5. Simulation

One can approach the simulation of BGNB processes in a variety of ways, using the representations presented in this work. All such methods depend on the ability to generate effectively sample paths of a gamma process, an issue explored extensively in the literature (see, e.g., [4] or [16]). Assuming we have a method of generating a standard gamma process, we shall use the representation given in Proposition 2.4 to simulate a path of BGNB process. First, we generate the following random samples, independently of one another:

Sample 1: A sample $G(t)$ of standard gamma process over a desired grid;
Sample 2: A sequence Γ_i of arrivals of standard Poisson process;
Sample 3: A sequence V_i of IID standard uniform variables;
Sample 4: A sequence E_i of IID standard exponential variables;
Sample 5: A sequence of independent Poisson variables with parameters $(1 - p^{V_i})E_i/p^{V_i}$.

Next, these samples are combined according to the representation described in Proposition 2.4, which results in discrete realization of a sample path of the BGNB process.

To illustrate the two-dimensional, real time process, in Figs. 1 and 2 we present two trajectories of one-dimensional coordinates, along with an image of the trajectory in the space of

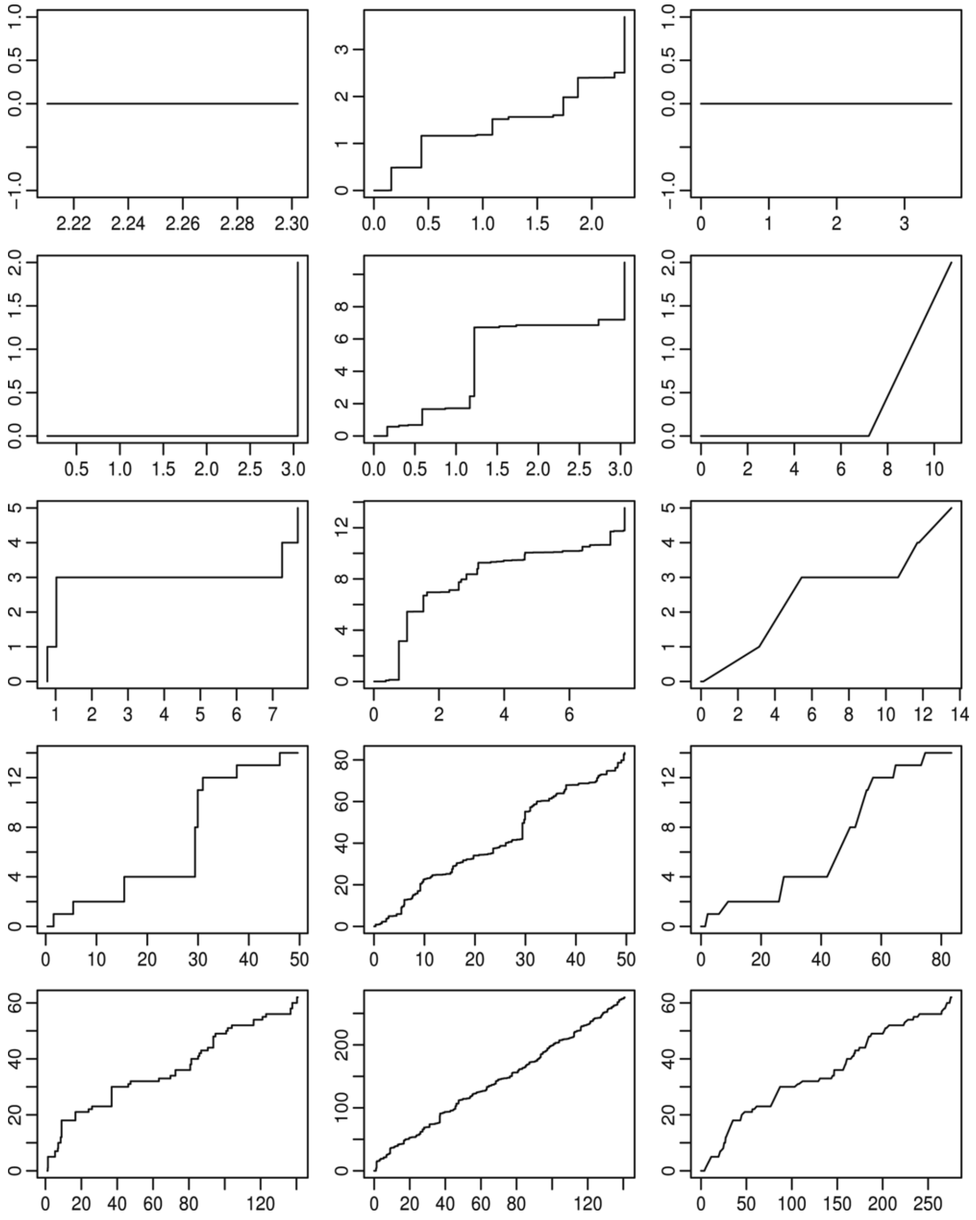

Fig. 1. Sample paths of BGNB process with $p = 0.5$, $\beta = 1$, and various time ranges as indicated on the graphs. The first and the second columns show the first and the second coordinates of the process, respectively, while the third column represents the image of the trajectory on the plane of values of the process.

values, i.e. $\mathbb{R}^2$-trajectory. The figures show typical sample paths of a BNBG process for various values of p and time ranges. The scale parameter of the first coordinate, β, is set to one. We observe that as p decreases, the role of the gamma process $G(t)$ diminishes, and the process $\mathbf{Y}_2$ (defined in Proposition 2.4) begins to dominate the behavior of the BNBG process.

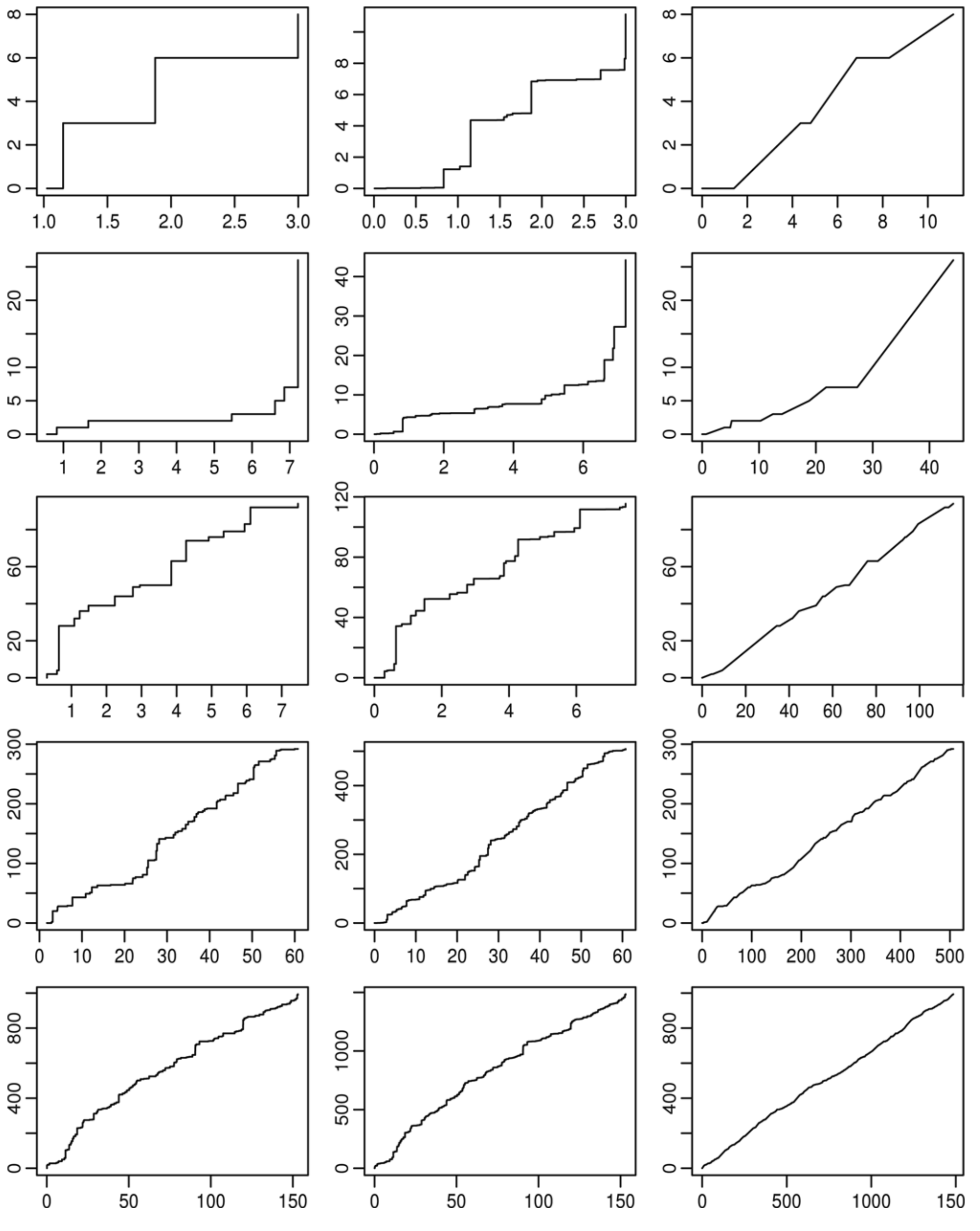

Fig. 2. Sample paths of BGNB process with $p = 0.1$, $\beta = 1$, and various time ranges as indicated on the graphs. The first and the second columns show the first and the second coordinates of the process, respectively, while the third column represents the image of the trajectory on the plane of values of the process.

6. Further extensions and remarks

As we mentioned in the introduction, the BGNB process studied in this paper is one of four possible models obtained by deleting/retaining either t or $G(t)$ (or both) on the right-hand-side of (4). Below we provide a brief description of the remaining three.

First, consider the "full" model $\{X(t), N(t), t \geq 0\}$ given by (4) when both t and $G(t)$ are retained. This is simply the BGNB process shifted by $(0, t)$. For each $t > 0$, the ChF is given by

$$\phi_{t,\beta,p}(r,s) = E\mathrm{e}^{\mathrm{i}\{rX(t)+sN(t)\}} = \left(\frac{p\beta \mathrm{e}^{\mathrm{i}s}}{\beta - ir - \beta(1-p)\mathrm{e}^{\mathrm{i}s}}\right)^t, \quad r, s \in \mathbb{R}, \tag{33}$$

while the PDF $g(x, y)$ is non-zero whenever $x > 0$ and $y = t + k$ with $k = 0, 1, 2, \ldots$, in which case it is given by the right-hand-side of (18). In particular, when t is a positive integer j, we obtain the distribution of the sum of j IID BEG random vectors, in which case the PDF takes the form

$$g_j(x,k) = \frac{\beta^k}{(j-1)!(n-j)!} x^{k-1} \mathrm{e}^{-\beta x} p^j (1-p)^{k-j}, \\ x > 0, k = j, j+1, j+2, \ldots, \tag{34}$$

and reduces to the BEG density when $j = 1$ (cf. equation (6) in [11]). Maximum likelihood estimation of the parameters $\beta > 0$, $p \in (0, 1)$, and $j \in \mathbb{N}$ based on a random sample $(X_1, Y_1), \ldots, (X_n, Y_n)$ from the model given by (34) is relatively straightforward. Here, the log-likelihood function is

$$L(\beta, p, j) = n\left\{\overline{Y}_n \log[\beta(1-p)] + j \log\frac{p}{1-p} - \log\Gamma(j) - \beta\overline{X}_n + D\right\}, \tag{35}$$

where D is the sample average of the quantities $\log[X_i^{Y_i-1}/(Y_i - j)!]$ and the positive integer j is bounded above by the $\min(Y_1, \ldots, Y_n)$. It is easy to see that the MLE of β is given by the ratio $\hat{\beta}_n = \overline{Y}_n/\overline{X}_n$, regardless of whether the other parameters are known or unknown. Further, for any given j the function (35) is maximized with respect to p by the quantity $\hat{p}_n(j) = j/\overline{Y}_n$, which can be regarded as the MLE of p when the parameter j is known. Note that when $j = 1$ the above MLEs reduce to those obtained by [11] for the BEG model, as expected. When j is unknown, its MLE, $\hat{j}_n$, can be found by maximizing the function $L(\hat{\beta}_n, \hat{p}_n(j), j)$, or equivalently, the function

$$h(j) = \overline{Y}_n \log\frac{\overline{Y}_n - j}{\overline{Y}_n} + j\log\frac{j}{\overline{Y}_n - j} - \log\Gamma(j) - \frac{1}{n}\sum_{i=1}^{n}\log[(Y_i - j)!] \tag{36}$$

over the set $j = 1, 2, \ldots, \min(Y_1, \ldots, Y_n)$, leading to the MLE of p, $\hat{p}_n = \hat{p}_n(\hat{j}_n) = \hat{j}_n/\overline{Y}_n$.

Next, consider the model $\{X(t), N(t), t \geq 0\}$ given by (4) when both t and $G(t)$ are deleted on the right-hand-side. The resulting stochastic representation,

$$\{(X(t), N(t)), t \geq 0\} \stackrel{d}{=} \left\{\left(\sum_{i=1}^{NB(t)} X_i, NB(t)\right), t \geq 0\right\}, \tag{37}$$

has an attractive interpretation. In insurance applications, $N(t)$ represents the frequency of claims in the time period $[0, t]$ while $X(t)$ is the corresponding total claim. The BGNB model has a similar interpretation as well; here, the quantity $G(t)$ in (11) can be thought of as a stochastic cost of running the business, which is on the average linear in t. The ChF corresponding to (37) can be written as

$$\psi_{t,\beta,p}(r,s) = \left(p + (1-p)\phi_{1,\beta,p}(r,s)\right)^t, \quad r, s \in \mathbb{R}, \tag{38}$$

where $\phi_{1,\beta,p}$ is the ChF (33) with $t = 1$, corresponding to the BEG model of [11]. Thus, for each $t > 0$ the marginal distribution of the Lévy process (37) is the t-power convolution of the mixture of an atom at zero (with probability p) and a BEG distribution (with probability $1-p$). In particular, the distribution of the first coordinate, $X(t)$, is the t-power convolution of the mixture of an atom at zero (with probability p) and an exponential distribution with parameter βp (with probability $1 - p$). For an integer value $t \geq 1$, we can write the ChF (38) in the form

$$\psi_{t,\beta,p}(r, s) = \sum_{j=0}^{t} \binom{t}{j} p^{t-j}(1-p)^j \phi_{j,\beta,p}(r, s), \quad r, s \in \mathbb{R}, \tag{39}$$

corresponding to an atom at (0, 0) with probability p^t and, with probability $1 - p^t$, a mixture of t convolutions of the BEG distributions, each given by the ChF (33) with $t = j$ and taken with probability $p_j = \binom{t}{j} p^{t-j}(1-p)^j/(1-p^t)$, $j = 1, 2, \ldots, t$. In view of (34), this interpretation allows us to write the PDF of the distribution given by (39) in the form

$$g_{t,\beta,p}(x, y) = \begin{cases} p^t \quad \text{for } x = 0, y = 0 \\ \dfrac{\beta^y}{\Gamma(y)} x^{y-1} \mathrm{e}^{-\beta x} p^t (1-p)^y \displaystyle\sum_{j=1}^{\min(t,y)} \binom{t}{j}\binom{y-1}{j-1} \\ \quad \text{for } x > 0, y = 1, 2, 3 \ldots \\ 0 \quad \text{otherwise.} \end{cases} \tag{40}$$

Note that while the joint distribution is rather complicated, the conditional distributions of $X(t)$ given $N(t) = n$ are quite simple. Indeed, for $n = 0$ we get a mass at zero with probability one, while for $n \geq 1$ we get a gamma distribution with shape parameter n and scale β. Let us now consider the problem of estimating the parameters $\beta > 0$, $p \in (0, 1)$, and $t \in \mathbb{N}$ based on a random sample $(X_1, Y_1), \ldots, (X_n, Y_n)$ from the model given by (40). Suppose that $n_0 \geq 0$ data points are double zeroes, while the rest of them can be put into k groups of size n_l each, where for $l = 1, 2, \ldots, k$ the data in the lth group consist of n_l pairs $(X_{l1}, Z_l), \ldots, (X_{ln_l}, Z_l)$ and $0 < Z_1 < Z_2 < \cdots < Z_k$ are k distinct non-zero values among all of the $\{Y_i\}$. In this notation, the log-likelihood function can be written as

$$L(\beta, p, t) = n\left\{\overline{Y}_n \log[\beta(1-p)] + t \log p - \beta \overline{X}_n + Q(t) + C\right\}, \tag{41}$$

where

$$Q(t) = \frac{1}{n} \sum_{l=1}^{k} \sum_{j=1}^{n_l} \log\left\{\binom{t}{j}\binom{Z_l - 1}{j-1}\right\} \tag{42}$$

depends only on the parameter $t \in \mathbb{N}$ while

$$C = \frac{1}{n} \sum_{l=1}^{k} \left\{(Z_l - 1) \sum_{i=1}^{n_l} \log X_{li} - n_l \log \Gamma(Z_l)\right\}$$

is parameter-free. Here, the MLE of β is again given by the ratio $\hat{\beta}_n = \overline{Y}_n/\overline{X}_n$, regardless of whether the other parameters are given or not. Further, for any fixed t the function (41) is maximized with respect to p by the quantity $\hat{p}_n(t) = t/(t + \overline{Y}_n)$, which can be viewed as the MLE of p when the parameter t is known. Note that the same MLE of p arises in the BGNB model (see (27)). When t is unknown, the MLE $\hat{t}_n$ can be found by maximizing the function

$L(\hat{\beta}_n, \hat{p}_n(t), t)$ over the set of integers $t \in \mathbb{N}$. This is equivalent to maximizing the function

$$v(t) = t \log \frac{t}{t + \overline{Y}_n} + Q(t) \tag{43}$$

and requires a numerical search. Once $\hat{t}_n$ is available, we can compute $\hat{p}_n = \hat{t}_n/(\hat{t}_n + \overline{Y}_n)$.

Finally, if in (4) we retain t and delete $G(t)$, we obtain another Lévy process given by the ChF

$$\psi_{t,\beta,p}(r, s)\mathrm{e}^{\mathrm{i}st}, \quad r, s \in \mathbb{R}, \tag{44}$$

with $\psi_{t,\beta,p}$ as in (38). Since this is simply the process discussed above shifted by $(0, t)$, it can be handled in a similar way. In particular, for each $t > 0$ the first coordinate is the t-power convolution of the mixture of an atom at zero (with probability p) and an exponential distribution with parameter βp (with probability $1 - p$). On the other hand, the second coordinate takes on the values $t + k$, $k = 0, 1, 2 \ldots$, with probabilities given by the right-hand-side of (16). While the joint distribution of $X(t)$ and $N(t)$ is again rather complicated, the conditional distribution of $X(t)$ given $N(t) = k + t$ is simply a mass at zero with probability one if $k = 0$ and gamma distribution with shape parameter k and scale β when $k \geq 1$.

7. Proofs

Proof of Proposition 2.2. Since both processes in (12) are Lévy, it is enough to compare their marginal distributions, say when $t = 1$. Conditioning on the variable $G(1)$, we obtain the following expression for the ChF of $(G(1), N(G(1)))$:

$$\begin{aligned}\mathbb{E}\mathrm{e}^{\mathrm{i}\{rG(1)+sN(G(1))\}} &= \mathbb{E}\left(\mathbb{E}\left\{\mathrm{e}^{\mathrm{i}\{rG(1)+sN(G(1))\}}|G(1)\right\}\right)\\ &= \int_0^\infty \mathbb{E}\mathrm{e}^{\mathrm{i}\{rx+sN(x)\}}\gamma \mathrm{e}^{-\gamma x}\mathrm{d}x.\end{aligned} \tag{45}$$

Since $\mathbb{E}\mathrm{e}^{\mathrm{i}sN(x)} = \exp\{\lambda x(\mathrm{e}^{\mathrm{i}s} - 1)\}$, elementary integration shows that, when $\gamma = p\beta$ and $\lambda = \beta(1 - p)$, the above ChF coincides with (6). Since the latter is the ChF of $(X(t), NB(t))$, the result follows. □

Proof of Proposition 2.3. Since the subordinated bivariate process (13) is a Lévy process, we proceed by showing that its marginal distributions coincide with those of a BGNB process. Conditioning on the variable $\widetilde{NB}_q(t)$ we express the relevant ChF as follows

$$\begin{aligned}&\mathbb{E}\mathrm{e}^{\mathrm{i}\{rX(\widetilde{NB}_q(t)+t)+sNB(\widetilde{NB}_q(t)+t)\}} = \mathbb{E}\left(\mathbb{E}\left\{\mathrm{e}^{\mathrm{i}\{rX(\widetilde{NB}_q(t)+t)+sNB(\widetilde{NB}_q(t)+t)\}}|\widetilde{NB}_q(t)\right\}\right)\\ &\quad= \sum_{k=0}^\infty \mathbb{E}\mathrm{e}^{\mathrm{i}\{rX(k+t)+sNB(k+t)\}}\mathbb{P}(\widetilde{NB}_q(t) = k) = \psi_{p,\beta}(r, s)G_q\left(\psi_{p,\beta}(r, s)\right),\end{aligned}$$

where $\psi_{p,\beta}(r, s)$ is the ChF (6) of the BGNB process $(X(t), NB(t))$ and

$$G_q(z) = \left(\frac{q}{1 - (1 - q)z}\right)^t$$

is the generating function corresponding to the NB variable $\widetilde{NB}_q(t)$. Straightforward algebra shows that the above ChF is of the form (6) with p and β replaced by p^* and β^* given in (14). This concludes the proof. □

For the proof of Proposition 2.4 we need the following interesting interpretation of the logarithmic series distribution.

Lemma 7.1. *Let the random variable $J = G_P$ be obtained from a geometric random variable through an independent randomization of its parameter, i.e. its parameter is a random variable $P \in [0, 1]$. Then J can be represented in distribution as*

$$J = 1 + \left[\frac{-W}{\log(1-P)}\right],$$

where W is exponentially distributed and independent of the randomly scattered P, and $[\cdot]$ is the integer part. In particular, if $P = p^U$, where U is a standard uniform variable, then J is distributed according to the logarithmic series distribution given by the PDF

$$\mathbb{P}(J_j = k) = -\frac{(1-p)^k}{k \log p}, \quad k \in \mathbb{N}.$$

Proof. The proof follows easily from the fact that a geometric random variable can be represented as $G_p = 1 + [-W/\log(1-p)]$, combined with the representation of the logarithmic series distribution presented in [7]. $\square$

Proof of Proposition 2.4. First, notice that by Proposition 2.1 we have (in distribution)

$$\mathbf{Y}(t) = (G(t), 0) + (G_1(NB(t)), NB(t)), \tag{46}$$

where $G(t)$, $G_1(t)$ are independent standard gamma processes independent of $NB(t)$. Indeed, both sides of (46) are Lévy processes, so it is enough to check the equality of distributions for $t = 1$, which is straightforward. Using the above lemma along with the representation of the negative binomial process as the Poisson process compounded by the logarithmic series distribution, we have

$$NB(t) = \sum_{i=1}^{\infty} J_i \mathbf{1}_{[\Gamma_i,\infty)}(t),$$

where $J_i = 1 + \left[-W_i/\log(1-p^{V_i})\right]$ (see, e.g., [12]). Thus, the process $G_1(NB(t))$ will have the jumps at the same time points Γ_i as the process $NB(t)$. Moreover, these jumps will be increments of Gamma process evaluated at the (integer) values of the jumps of $NB(t)$. Conditionally on $V_i = v$, the ith jump of $NB(t)$ is geometrically distributed with the parameter p^v. Given $V_i = v$, by the stochastic self-similarity (see also (9)) we obtain

$$G_1(NB(\Gamma_i)) - G_1(NB(\Gamma_{i-1})) \stackrel{d}{=} E_1 + \cdots + E_{J_i} \stackrel{d}{=} p^{-v} E_i.$$

This concludes the argument. $\square$

Proof of Proposition 4.1. First, note that the function R given by (28) is continuous and differentiable on $[0, \infty)$ and $(0, \infty)$, respectively, with the derivative

$$\frac{\mathrm{d}R(\alpha)}{\mathrm{d}\alpha} = \overline{(\log X)_n} - \log \overline{X}_n + \log\alpha - \psi(\alpha). \tag{47}$$

Second, using the well-known integral representation of the Digamma function (see, e.g., equation 6.3.21 in [1], p. 259), write

$$\log\alpha - \psi(\alpha) = \frac{1}{2\alpha} + 2\int_0^{\infty} \frac{u\,\mathrm{d}u}{(u^2+\alpha^2)(\mathrm{e}^{2\pi u}-1)}, \quad 0 < \alpha < \infty. \tag{48}$$

Observe that the function on the right-hand-side of (48) is positive and monotonically decreasing from infinity to zero as α varies between zero and infinity. On the other hand, by the concavity of the logarithmic function, we have

$$\overline{(\log X)}_n \le \log \overline{X}_n, \tag{49}$$

with strict inequality whenever not all of the $\{X_j\}$ are the same. We conclude that, under the latter condition, the derivative (48) is monotonically decreasing from infinity to $\overline{(\log X)}_n - \log \overline{X}_n < 0$ as α varies between zero and infinity. Consequently, the function R attains its maximum value at a unique solution of Eq. (29). Finally, when all the $\{X_j\}$ coincide (which occurs with probability zero), then $\overline{(\log X)}_n = \log \overline{X}_n$ so that the derivative of R is positive and the function R is strictly increasing. The result follows. □

To prove Proposition 4.2 we need the following technical lemma.

Lemma 7.2. *Let $W = (X, \log X, Y)$, where $(X, Y) \sim \mathcal{BGNB}(\alpha, \beta, p)$. Then the mean vector and the covariance matrix of W are*

$$\mathbb{E}W = \left(\frac{\alpha}{\beta p}, \psi(\alpha) - \log(\beta p), \frac{\alpha(1-p)}{p}\right) \tag{50}$$

and

$$\Sigma_W = \frac{1}{\beta^2 p^2} \cdot \begin{bmatrix} \alpha & \beta p & \alpha\beta(1-p) \\ \beta p & \psi'(\alpha)\beta^2 p^2 & \beta^2 p(1-p) \\ \alpha\beta(1-p) & \beta^2 p(1-p) & \alpha\beta^2(1-p) \end{bmatrix}, \tag{51}$$

respectively, where $\psi(\alpha)$ is the Digamma function and $\psi'(\alpha)$ is the Trigamma function.

Proof. The means and variances of X and Y follow from Corollary 3.1. To compute the mean of $\log X$, recall that X has a gamma distribution with shape parameter α and scale βp and use the identity ([9], p. 604)

$$\int_0^\infty x^{v-1} \mathrm{e}^{-\mu x} \log x \mathrm{d}x = \frac{\Gamma(v)}{\mu^v}(\psi(v) - \log \mu) \tag{52}$$

with $v = \alpha$ and $\mu = \beta p$. The variance of $\log X$ is computed similarly via the relation ([9], p. 607)

$$\int_0^\infty x^{v-1} \mathrm{e}^{-\mu x} [\log x]^2 \mathrm{d}x = \frac{\Gamma(v)}{\mu^v}\{[\psi(v) - \log \mu]^2 - \zeta(2, v)\}, \tag{53}$$

where

$$\zeta(z, v) = \sum_{n=0}^\infty \frac{1}{(v+n)^z}$$

is Riemann's zeta function. Applying (53) with the same values of v and μ as above and using the identity $\zeta(2, \alpha) = \psi'(\alpha)$ we obtain the required expression for the variance. The computation of $\mathbb{E}(x \log x)$ is quite similar, where Eq. (52) is used again (this time with $v = \alpha + 1$ and $\mu = \beta p$) along with the relation $\psi(\alpha + 1) = \psi(\alpha) + 1/\alpha$. Finally, noting that the infinite series in the expression

$$\mathbb{E}(Y \log X) = \int_0^\infty \frac{\log x}{\Gamma(\alpha)} \beta^\alpha x^{\alpha-1} \mathrm{e}^{-\beta x} p^\alpha \sum_{k=0}^\infty k \frac{\beta^k}{k!} x^k (1-p)^k \mathrm{d}x \tag{54}$$

simplifies to $\beta x(1-p)\exp\{\beta x(1-p)\}$ and using (52) one last time (again with $\upsilon = \alpha + 1$ and $\mu = \beta p$) we arrive at the required expression for the covariance of $\log X$ and Y. This concludes the proof. □

Proof of Proposition 4.2. Since the MLE $\hat{\alpha}_n$ is a unique solution of (29), it can be written as

$$\hat{\alpha}_n = H_1(\overline{X}_n, \overline{(\log X)}_n),$$

where $H_1(\cdot,\cdot)$ is a continuous and differentiable function satisfying the equation

$$F(y_1, y_2, H_1(y_1, y_2)) = 0 \tag{55}$$

with

$$F(y_1, y_2, y_3) = y_2 - \log y_1 + \log y_3 - \psi(y_3). \tag{56}$$

This allows us to express the MLEs as

$$(\hat{\alpha}_n, \hat{\beta}_n, \hat{p}_n) = H(\overline{X}_n, \overline{(\log X)}_n, \overline{Y}_n), \tag{57}$$

where $H(y_1, y_2, y_3) = (H_1(y_1, y_2), H_2(y_1, y_2, y_3), H_3(y_1, y_2, y_3))$ with $H_1(\cdot,\cdot)$ as above and

$$H_2(y_1, y_2, y_3) = \frac{H_1(y_1, y_2) + y_3}{y_1}, \qquad H_3(y_1, y_2, y_3) = \frac{H_1(y_1, y_2)}{H_1(y_1, y_2) + y_3}. \tag{58}$$

The consistency and asymptotic normality of the MLEs now follow from standard large sample theory (see, e.g., [15]). Indeed, by Lemma 7.2 and the law of large numbers applied to the sequence $W_i = (X_i, \log X_i, Y_i)$ we conclude that the sample mean vector $(\overline{X}_n, \overline{(\log X)}_n, \overline{Y}_n)$ converges in probability to the mean vector $\mathbb{E}W = (\mu_1, \mu_2, \mu_3)$ given by (50). Thus, by the continuity of H we obtain the convergence of the MLEs (57) to the quantity

$$H(\mu_1, \mu_2, \mu_3) = (\alpha, \beta, p). \tag{59}$$

To verify (59), note that the triple $y_1 = \mu_1$, $y_2 = \mu_2$, $y_3 = \alpha$ satisfies the equation $F(y_1, y_2, y_3) = 0$, so that

$$H_1(\mu_1, \mu_2) = \alpha. \tag{60}$$

In turn, using the above along with the expressions (58) we obtain

$$H_2(\mu_1, \mu_2, \mu_3) = \frac{\alpha + \alpha(1-p)/p}{\alpha/(\beta p)} = \beta \tag{61}$$

and

$$H_3(\mu_1, \mu_2, \mu_3) = \frac{\alpha}{\alpha + \alpha(1-p)/p} = p. \tag{62}$$

This concludes the consistency (i). Similarly, we establish the asymptotic normality (ii) of the MLEs. By the classical multivariate central limit theorem, we have the convergence in distribution

$$\sqrt{n}[(\overline{X}_n, \overline{(\log X)}_n, \overline{Y}_n) - (\mu_1, \mu_2, \mu_3)] \xrightarrow{d} \mathrm{N}(0, \Sigma_W),$$

where the right-hand-side denotes the trivariate normal distribution with mean vector zero and variance–covariance matrix Σ_W given by (51). Now, by standard large sample theory (see, e.g., [15]), as $n \to \infty$, the variables

$$\sqrt{n}[H(\overline{X}_n, \overline{(\log X)}_n, \overline{Y}_n) - H(\mu_1, \mu_2, \mu_3)] = \sqrt{n}[(\hat{\alpha}_n, \hat{\beta}_n, \hat{p}_n) - (\alpha, \beta, p)]$$

converge in distribution to a trivariate normal distribution with mean vector zero and variance–covariance matrix

$$\Omega = D\Sigma_W D', \tag{63}$$

where

$$D = \left[\left. \frac{\partial H_i}{\partial y_j} \right|_{(y_1,y_2,y_3)=(\mu_1,\mu_2,\mu_3)} \right]_{i,j=1}^{3}$$

is the matrix of partial derivatives of the vector-valued function H. Since the function H_1 satisfies Eq. (55), we have

$$\frac{\partial}{\partial y_i} H_1(y_1, y_2) = \frac{\frac{\partial}{\partial y_i} F(y_1, y_2, y_3)|_{y_3=H_1(y_1,y_2)}}{\frac{\partial}{\partial y_3} F(y_1, y_2, y_3)|_{y_3=H_1(y_1,y_2)}}, \quad i = 1, 2, \tag{64}$$

with $F(\cdot, \cdot, \cdot)$ given by (56). Utilizing the above along with relations (56), (58) and (60)–(62), after rather lengthy albeit straightforward calculations we arrive at

$$D = \frac{1}{\alpha C} \begin{bmatrix} \alpha\beta p & -\alpha^2 & 0 \\ \beta^2 p^2 (1 - C/p) & -\alpha\beta p & \beta p C \\ \beta p^2 (1-p) & -\alpha p(1-p) & -p^2 C \end{bmatrix},$$

with C defined in Proposition 4.2. This combined with (51) and (63) produces the asymptotic variance–covariance matrix (32). Finally, inverting the Fisher information matrix (31) we obtain (32), proving the asymptotic efficiency (iii). □

Acknowledgements

We thank the referee for useful suggestions. The research of the first author was partially supported by NSF grant ATM-0503722. The research of the second author was partially supported by NSF grants ATM-0236898 and ATM-0503722.

References

[1] M. Abramowitz, I.A. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, ninth printing, Dover Publications, New York, 1972.

[2] F. Biondi, T.J. Kozubowski, A.K. Panorska, Stochastic modeling of regime shifts, Climate Res. 23 (2002) 23–30.

[3] F. Biondi, T.J. Kozubowski, A.K. Panorska, A new model for quantifying climate episodes, Internat. J. Climatology 25 (2005) 1253–1264.

[4] L. Bondesson, On simulation from infinitely divisible distributions, Adv. Appl. Probab. 14 (1982) 855–869.

[5] R. Cont, P. Tankov, Financial Modelling with Jump Processes, Chapman & Hall/CRC, Boca Raton, 2004.

[6] E.R. Cook, P.J. Krusic, The North American drought atlas, EOS Transactions of the American Geophysical Union 84 (2003) Abstract GC52A-01.

[7] L. Devroye, Non-Uniform Random Variate Generation, Springer-Verlag, New York, 1986.

[8] T.S. Ferguson, M.J. Klass, A representation of independent increment processes without Gaussian components, Ann. Math. Statist. 43 (1972) 1634–1643.

[9] I.S. Gradshteyn, I.M. Ryzhik, in: A. Jeffrey (Ed.), Table of Integrals, Series, and Products, 5th edition, Academic Press, San Diego, 1994.

[10] T.J. Kozubowski, M.M. Meerschaert, K. Podgórski, Fractional Laplace motion, Adv. Appl. Probab. 38 (2006) 451–464.

[11] T.J. Kozubowski, A.K. Panorska, A mixed bivariate distribution with exponential and geometric marginals, J. Statist. Plann. Inference 134 (2005) 501–520.

[12] T.J. Kozubowski, K. Podgórski, Distributional properties of the negative binomial Lévy process, 2005, preprint.

[13] T.J. Kozubowski, K. Podgórski, Invariance properties of the negative binomial Lévy process and stochastic self-similarity, Intern. Math. Forum 2 (30) (2007) 1457–1468.

[14] L. Mathier, L. Perreault, B. Bobée, F. Ashkar, The use of geometric and gamma-related distributions for frequency analysis of water deficit, Stochastic Hydrology Hydraulics 6 (1992) 239–254.

[15] C.R. Rao, Linear Statistical Inference and its Applications, Wiley, New York, 1973.

[16] J. Rosiński, Series representations for Lévy processes from the perspective of point processes, in: O.E. Barndorff-Nielsen, T. Mikosch, S.I. Resnick (Eds.), Lévy Processes: Theory and Applications, Birkhaüser, Boston, 2001, pp. 401–415.

[17] P. Todorovic, D.A. Woolhiser, Stochastic structure of the local pattern of precipitation, in: H.W. Shen (Ed.), Stochastic Approaches to Water Resources, Vol. 2, Colorado State University, Fort Collins, Colorado, 1976.

[18] E. Zelenhasić, A. Salvai, A method of streamflow drought analysis, Water Resour. Res. 23 (1) (1987) 156–168.